

1

导论

材料、能源和信息是人类社会赖以生存的基础。与之对应,在科学认知上人们经过了从认知物质、能量到认知信息的升华。很难给信息(information)下一个明确的定义,信息科学的3位先驱者作了不同的描述。控制论之父 Norbert Wiener(1894—1964 年,美国)指出,信息是人们在适应外部世界并使这种适应反作用于世界的过程中同外部交换内容的总称。通信理论创始人 Claude E. Shannon(1916—2001 年,美国)给出信息的数学表达为一个事件发生概率的对数的负值。数学物理学家 Louis Marcel Brillouin(1854—1948 年,法国)则认为信息是接收前后可能答案的关系的函数。

人的信息摄取能力是非常强大和相当高级的,人类通过眼、耳、皮肤(以四肢为主)、口以及鼻等感觉器官从外界获取信息。有资料表明:在人所获得的全部信息来源中,视觉约占 60%,听觉占 20%,肤觉占 15%,味觉占 3%,嗅觉占 2%。如果再考虑借助各种传感器感知不可见光的图像(例如超声、微波、红外、紫外、X 光、伽玛和中子等粒子成像),以及利用电子与计算机技术将其他信息转换成可视的图像,例如各类图表以及视窗操作系统、可视化计算、可视化编程和虚拟现实等人机交互技术,可以毫不夸张地说,人类获取的信息 70%~80% 是以图像形式提供的。

图像技术在古代作为印刷技术,18 世纪后又作为光学技术(例如照相和电影)发展起来。第 2 次世界大战后随着电视的出现和普及,使得仅在教育光学领域中处理的图像与电子工程紧密联系起来,成为电子学的重要领域。特别是在近年来由电视、通信和计算机三个领域相互融合产生的多媒体(multimedium)技术已经成为了现代信息高速公路的核心。

总之,通过可视能够有效地组织与提取有用信息,利用可视化可

以有效地表示信息。由于图像科学与技术的普遍性和有效性,确立了它在先进的信息处理领域内居中心的地位。

1.1 信息媒体与图像信息

信息有三方面特征:语法特征(包括信息的语法、存储和传递的描述)、语义特征(信息的含义)和语用特征(信息总是为一定的目的服务)。Shannon 定义信息的数学表达方式主要是表达了信息的语法特征。Brillouin 的定义指出信息将减少不确定性,是语义特征。Wiener 则强调了信息的语用特征,这在控制论以及与信息源的反馈过程中反映得非常明显。

如同空间要依赖于物质而存在那样,必须具有信息的载体(或称媒体),才存在信息。现代的社会是信息化的社会,信息在生活中不可缺少和普遍存在。如果说“视”是人类最有效和最重要的信息获取方式(正如常说的百闻不如一见),“听”与“说”则是人类最方便的信息交流方式,那么,多媒体技术通过使用计算机把各种电子媒体集成起来,完成视、听、说、记忆、通信等功能,成为了人类交流信息的重要媒体。根据国际电信联盟电信标准化部门(International Telecommunications Union—Telecommunication Sector, ITU-T)的相关文件,信息媒体可以有以下 5 种定义。

(1) 感觉媒体(perception medium):能直接作用于人的感官,使人产生感觉的媒体。包括人类的语言、音乐和自然界的各种声音、活动图像、静止图像、图形、曲线、动画、文本等。

(2) 表示媒体(representation medium):为传输感觉而使用的中间手段,以便能更有效地将感觉从一地传到另一地。表示媒体包括说明交换信息的类型,定义信息的特征,一般以编码的形式描述,如声音编码、图像编码、文本编码等。

(3) 显示媒体(presentation medium):用于通信中电信号和感觉媒体之间转换所用的媒体。显示媒体主要包括获取和显示信息的输入、输出设备两种。常用的输入设备有键盘、鼠标、摄像机、扫描仪、光笔和话筒等,输出设备有显示器、打印机、音箱等。

(4) 存储媒体(storage medium):用于存储表示媒体,以便使本机随时调用或供其他终端远程调用。存储媒体主要是存储数据的物理设备,如磁带、软盘、硬盘、光盘等。

(5) 传输媒体(transmission medium):用来表示媒体从一地传输到另一地的物理实体。传输媒体主要由传输数据的物理设备组成,如双绞线、铜轴电缆、光纤和无线电波等。

图像可以简单定义为在二元(或三元)画面上呈现出来的视觉信息,它是周围世界能量场的一种映像,是由空间三维坐标、能量波长分布和时间为变量的多维函数。这些由空间信息形成的图像在人的视觉系统和大脑中引起感觉、知觉、情绪等各种程度的变化。但是包括人眼在内的大部分成像系统得到的仅是二维平面图像,而且对能量波有积累效应,

因此往往不必精确重现景物辐射的能量波长分布,这就使得图像工程中要实现摄像以及图像的重现才有可能。

图像信息的内容大体上有两种表示方法:把物体作为图像来表示,称为直觉性符号(presentational symbol);以及把符号化了的东西(例如文字、图画等)作为图像来表示,叫做推论性符号(discursive symbol)。图像或影像(image)与图画(picture)的区别在于前者是逼真地重现,后者是相似。实际上,实物的信息量很大,要完全一模一样重现不仅技术难度大,也没有必要。因为人类视觉本身的限制会丢失景物中过分精细的信息,而且观察者往往是按照应用要求和本人愿望来选择对图像的某些部分有所重视和偏爱,因此,图像信息系统强调的是图像的有用性并非全部内容,它采用的技术手段包括应用光学、光化学、机械和电子等各种技术。

图像一般是作为具有一定的时间和具有空间频率能谱的信号来处理的,时间和空间频谱定义是在统计方面处理图像的。图像信号形式可以分成模拟图像和数字图像。数字图像是数据量大的二维结构,它实时传输时的数据码率高,对图像系统有很高要求。如果按照数据结构的观点以及从光谱、空间与时间轴等方面考虑,对图像 $f(x,y)$ 进行分类,可以分为如下几类。

- 二值图像: $f(x,y)=0,1$,例如文字、线图形等。
- 灰度图像: $0 \leq f(x,y) \leq 2^n - 1$,例如黑白照片等。
- 彩色图像:通常用三基色分量图像或者亮度与色度分量图像联合表示, $\{f_i(x,y)\}$, $i=R,G,B$ 或 Y,U,V 或 Y,I,Q 等。
- 多光谱图像: $\{f_i(x,y)\}, i=1, \dots, m$,例如遥感图像等。

如果考虑图像随时间变化的情况,又可以分为没有动作印象的静止图像和活动图像(时间序列图像): $\{f_i(x,y)\}, t=t_1, \dots, t_n$ 。

另外,虽然常见的二维平面图像也包含着景物深度 z 的信息,但是栩栩如生的立体图像(stereoscopic image)更会受到人们的青睐。立体图像一般有两种:一种是由左、右两眼各自观看同一景物的两幅视图,视觉系统使之融合产生立体图像,称为双目视立体像。另一种是在观众眼前呈现三维的全息立体图像,称为三维图像(3-dimensional image)。

1.2 图像信息系统

人类需要掌握技术是因为人类本身的感觉机能和行动机能不能满足人类的需求,其中对于视觉系统有两种:第一是想会见远方亲人的欲望,从而发展了望远镜和电视;第二是想永远记住的欲望,为此产生了照相或印刷技术。这些可以看作是人类感觉机能的空间延长和时间延长,就图像科学而言就是图像传输技术和图像记录技术。此外,正在发展扩大视觉识别机能的技术,例如把超出视觉能力的红外线和紫外线图像转换成可视像;

把细节模糊的图像处理得更清晰等,这种技术称为图像处理技术(image processing)。

图像处理技术最初是和照相以及电影技术一起发展起来的,是应用光学机械和光化学技术处理自然可见光图像的空间信息,它属于同时制(simultaneous system)。后来它们与电子技术结合,产生了电子照相、电子印刷等新的技术领域。被誉为视觉革命的电视,它的发明使得仅在教育应用光学领域中处理的图像与电子工程紧密联系起来,而电子学的发展给图像的传送、存储、处理等各技术领域带来了巨大的变革。长期以来,在电子学领域内处理的信息和信号都是时间函数。图像电子学利用扫描方法(scanning)把图像的空间信息变换成便于处理的时间信息,它属于顺序制(sequential system)。另外,半导体技术的发展、各种新的电光和光电转换器件的产生,使得摄像与显像设备日新月异,因此图像在光电子领域中占有很大的比重。20世纪50年代后期激光技术的出现,又使电子工程学中的工作波长从微波波段扩展到光波波段。这就促使光学领域与图像电子学领域更加接近和融合,为图像信息的记录、显示以及传输增添了新的发展空间。

一般最常见的图像信息系统是首先把目标物体的信息(多数是光学像,但不局限于光学像)变成容易传输、记录和处理的电信号,然后再经过传输、记录和处理后恢复成视觉信息(通常是光学像)。图1.1是接收视觉信息的通用图像信息系统的概念模型,该模型符合Shannon通信模型的形式(括号内是Shannon通信模型的提法)。

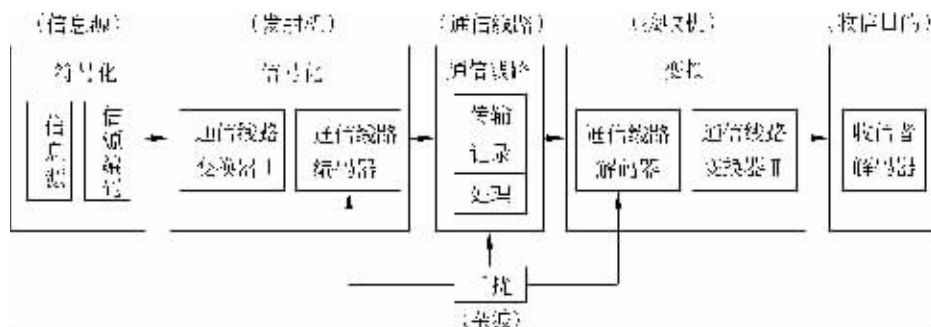


图 1.1 图像信息系统的概念模型(宫川洋、渡部毅)

其工作原理简述如下: 首先把想要告诉对方的视觉信息客观化(符号化),然后把符号变换成物理存在的信号并按后续的通信线路便于处理的形式编码(信号化),例如用摄像管把光像转换成电信号等。图像信号再经过广义通信线路(诸如传输、记录、处理信号的装置)处理后送往接收端。在接收端,通过解码把信号作为可视的空间图像来显示,进而引起接受者的感觉和知觉。最后收信者(信息论称其为信宿)把引起的感觉、知觉和过去的学习、记忆相对照,从而认识信息源的内容或者刺激情绪。另外,在光与电的转换以及信道传输过程中会引入干扰和噪声。通常,图像的噪声大致可分为3类:感光照片上的颗粒噪声(可以用高斯白噪声作为模型)、光电转换过程中产生的光电子噪声(光照较弱

时,可以用具有泊松分布的随机变量作为模型,在光照较强时,泊松分布趋向高斯分布),以及电子设备产生的热电子噪声(一般常用零均值的高斯白噪声作为模型)。根据信息论的信息不增性,在数据处理过程中总会失掉一些信息,而且用任何处理手段也不可能再恢复。同时,噪声也会增加对信源符号的疑义度(平均不确定度),另一方面人类视觉系统对信息也有选择过滤作用。因此,可以充分利用人眼的视觉特性以及针对不同的应用需求来选择具有不同性能要求的图像信息系统,追求最佳的系统性能价格比。

1.3 图像技术的发展和涉及的关键技术

图像技术经过了印刷术、摄影术、光电子技术与计算机技术结合的发展阶段。下面仅列举出 20 世纪以前一些具有标志性的发展事件:

1837 年法国人 L. M. Daguerre(1787—1851 年)发明银板照相。

1884 年 11 月俄裔德国科学家 Paul-Nipkow(1886—1940 年)申报了电望远镜发明专利(世界电视史上的第一个专利),这就是著名的 Nipkow 圆盘。

1895 年法国 Lumière 兄弟 Louis(1862—1954 年)与 Auguste(1864—1948 年)在巴黎“大咖啡馆”放映《火车到站》,这标志着电影的诞生。

1895 年,俄国人 A. C. Popov 和意大利人 Guglielmo Marconi 分别在俄罗斯和意大利独立地实现了第一次无线电传输。

1897 年德国人 K. F. Braun(1850—1918 年)设计 CRT 显像管方案。

1925 年美籍俄国人 V. K. Zworykin(1889—1982 年)发明积累式摄像方式。

1926 年 1 月英国人 J. L. Baird(1888—1946 年)发明机械扫描电视。

1936 年英国广播公司(BBC)在伦敦开播 405 行每秒 25 帧的黑白电视广播。

1945 年年底美国宾夕法尼亚大学的莫尔学院和美国陆军阿伯丁弹道研究实验室研制出电子数值积分计算机(ENIAC),1946 年 2 月公开展示世界上第一台电子计算机。

1950 年美国第一台和麻省理工学院(MIT)的 Whirlwind I 型计算机联用的显示器产生了简单的画面,计算机图形学出现了曙光。

1953 年美国联邦通信委员会(FCC)确定 NTSC 制,这是第一个广播的彩色电视制式。

1960 年美国第一台红宝石激光器起振,激光技术为图像工程带来巨大的进步。

1965 年美国人 Ted Nelson 提出计算机处理文本的超文本(hypertext)方法。

1968 年美国国防部高级研究计划署(DARPA)将 Web 网络项目交给加州大学洛杉矶分校,1969 年研制出 ARPANET 分组(packet)交换计算机网络。1972 年 Xerox 公司开发出以太网(Ethernet)。

1972 年英国独立广播公司(IBA)研制出第一台数字式电视制式转换设备。

1974 年美国首次成功进行 2.6GHz 和 800MHz 定点接收的卫星广播。

1976 年美国 MIT 计算机结构组 (Architecture Machine Group) 向 DARPA 提出多媒体 (Multiple Media) 的建议。

1985 年美国 Microsoft 公司推出了多用户的图形操作系统 Windows。

1991 年日本通过 BS-2 卫星每天 8 小时播出 MUSE 高清晰度模拟电视广播。

1993 年美国率先提出传输速度为 Gbps 的信息高速公路计划。

1994 年在瑞士日内瓦举行的第一届国际互联网 (WWW) 大会上提出要有用于互联网的超级链接语言, 于是诞生了虚拟现实造型语言 (Virtual Reality Modeling Language, VRML)。

1997 年美国 FCC 向 ABC、CBS、NBC 和 FOX 颁发数字电视试播执照, 并宣布到 2006 年美国将全面取消模拟电视广播。

.....

图像技术特别是数字图像技术最具爆炸性的发展无疑是在 20 世纪 70 年代以后, 如果没有快速傅里叶变换 (FFT) 算法在 20 世纪 60 年代中期的普及, 现在许多图像变换和处理差不多还是光学领域的事。如果没有超大规模半导体集成电路 (VLSI) 设计和制造技术对图像系统构成的巨大推动、先进的计算机系统结构 (CA) 对实现图像处理算法的重要性以及日益重要的人工智能 (AI) 技术的发展, 就没有数字图像及其处理技术。因此, 相当长的一段时间内图像处理的发展主要与这三个关键技术的发展有关。

1. 数字图像的特点

二维结构的数字图像具有数据量大、处理实时性强, 传输数据码率高的特点, 这对图像系统的模数与数模转换器、存储器和处理器都有很高的要求。这里仅仅举一幅, 或称为一帧 (frame) 水平和垂直方向均为 512 像素, 每像素 8b 组成的中等分辨率的灰度图像来说, 它的数据量就是 $512 \times 512 \times 8 = 2\,097\,152\text{b} = 256\text{KB}$ 。如果它还是以 25 帧/秒运动的活动图像, 实时处理与传输一幅图像的时间只有 40ms, 因此, 其数据码率 $= 2\,097\,152 \times 25 = 52\,428\,800\text{bps}$ 。

2. 超大规模半导体集成电路技术

超大规模集成电路 (very large scale integrated circuit, VLSI) 主要从集成度、工作频率、功耗和成品率等几方面考虑。1965 年仙童 (Fairchild) 的设计工程师 Gordon Moore 发表的论文分析了从 1959—1965 年半导体工业发展的数据, 得出 Moore 定律: 集成电路上集成的晶体管数大约每 18~24 个月增长一倍。具体来说, 20 世纪 60—70 年代几乎每年翻 2 番, 以后每两年翻 1~2 番。集成度的提高依赖于减小线宽, 现阶段的产品可做到亚微米和纳米 (2005 年年底 Intel 推出采用 65nm 工艺的 Yonah 核心 Pentium M 处理器, Intel 研究结果表明大约 2018 年进入 16nm 制造工艺时代就可能达到物理极限)。集成电

路工作频率的提高也依赖于缩短内部连线(在片内 on-chip 通信比片外 off-chip 更可取),而更重要的是半导体材料的特性(主要是能带间隙和电子迁移率)。常用的硅材料工作频率限制在几个 GHz,而且它的导通电阻比较大,很费电。但是一些由两种 III—V 族或 II—VI 族元素制备的复合材料,它们的带隙(band gap)窄、电子迁移率(mobility)高,适于高频应用。其中主要有砷化镓(GaAs)、硅锗(SiGe)和磷化铟(InP),例如 IBM 发明的硅锗材料的工作频率可达到 100GHz。在基础研究上的开发包括:共振隧穿器件(RTD)、单电子晶体管(SET)、量子电子器件、分子电子器件、自旋电子器件等新原理的器件。通过降低芯片的工作电压可以压低信号波形跳变的幅度以减小互相干扰和降低功耗。例如从早期的 12V~15V MOS 工作电压、TTL 的 5V 到 CMOS 的 3V, 2.5V, 1.475V, 而 Dothan 1.7GHz 处理器的工作电压已为 0.956V~1.052V。特别是采用智能节电技术(例如 Intel 65nm 的 SRAM 采用空闲期切断电流的睡眠晶体管)十分有益于采用电池工作的移动设备。提高集成电路的成品率以降低价格要依靠故障容限技术(例如可以给存储单元增加备份,在测试过程中用电气或激光把不合格单元的连线割去)、冗余和自动恢复技术,以及增大晶体圆片(wafer)的尺寸等。目前最大的圆片直径:硅为 300mm(约 12in, 1in=25.4mm),硅锗为 200mm(约 8in),砷化镓为 150mm(约 6in),磷化铟为 100mm(约 4in),预计 2016 年左右 450mm(约 18in)硅片将投入生产。

另外,应该关注对图像处理系统的底层硬件开发有重要意义的可重构系统芯片(system on chip, SOC)以及圆片规模集成技术(wafer scale integration, WSI),将来 WSI 比传统的单基底 VLSI 在功能上提高 2~3 个数量级。例如一个 WSI 的 VLSI 器件可以包括一个智能的帧存储器和采用 CCD(charge coupled device) 技术的图像传感器等,它就能构成一个实用的图像处理系统。另一个发展方向是三维集成电路,其最初的目标是双层晶体管。

3. 计算机系统结构

计算机系统结构(computer architecture)根据其指令(I)和数据流(D)的单一性(S)或多重性(M)来分类,有 4 类: SISD、SIMD、MISD、MIMD。目前通用计算机仍然是 20 世纪 30—40 年代 Von Neumann 等人提出的存储程序,具有串行控制流程,并由通信单元分别连接处理单元和线性存储空间的结构,属于 SISD。虽然在某些 CPU(采用 MMX/SSE 技术)的内部实现了并行计算(例如有 4 个 16 位寄存器和其独立的数据通道,可以同时完成 4 个 16 位整数的加、减、乘运算)属于 SIMD,但是总体结构并没有变。可喜的是最近 CPU 芯片向着单芯多核(chip multiprocessor, CMP)发展来提高其性能,例如 Intel 2006 年第一季度推出 Dempsey 的 65nm 双内核 Xeon 芯片和 AMD 的双核 Opteron。此外,在数字信号处理领域还常常使用数字信号处理器(digital signal processor, DSP),它具有硬件乘法器和多个并行操作的功能单元、程序和数据分开的总线结构、经过专门优化的流水线指令操作等并行技术与专用的功能模块,特别适合滤波、卷积、编解码和 FFT 等

数字信号处理。但是,最适合图像处理的计算结构最好是映射到图像数据本身的二维并行处理,自从 1958 年 Unger 提出模式识别使用由二维处理器阵列组成的结构以来,在多个处理器的图像处理计算机方面,已经把精力集中到研制 SIMD。这期间有 20 世纪 70 年代中期的细胞逻辑图像处理机 CLIP,分布式阵列处理机 DAP 和巨型并行处理机 MPP,它们适合作高数据吞吐量的低层次的处理。而高性能的图像处理机采用 MIMD 结构,并装备了面向过程的并发操作的专用微处理器(例如 Inmos 公司的 1985 年 Transputer)。由于大部分的图像处理是把诸如特征提取这一类低层次的处理到特征和物体识别等高层次的运算分成了逐级处理功能,称之为处理锥这样一种金字塔结构。因此,强烈主张在低层次上采用高吞吐量的 SIMD 阵列,而高层次采用具有人工智能的 MIMD 结构来完成。另外,一直是制约计算机内部通信瓶颈的系统总线也在不断地发生革命性的变化,例如从个人计算机(personal computer,PC)二十余年的发展来看,从数据位宽 8 位、2Mbps 左右传输速率的 PC/XT 总线发展到 64 位并行总线 PCI-X 533(其共享带宽达到 4.2Gbps),以及正在发展的采用串行方式的 PCI Express 总线(在 PCI Express 标准中,以一个 Lane 信道实现的通信称为 $\times 1$ Link,PCI Express $\times 1$ 则表明有 1 组 Link,在双向传输的时候速度为 500Mbps,PCI Express 的最高规格 $\times 32$ 可以达到 16Gbps 的传输速度)。

值得指出,正在探讨利用生物芯片、神经网络芯片等来实现计算机发展的突破。但是在传统计算机的基础上大幅度提高性能必将遇到难以逾越的障碍,根本的还是要从基本原理上寻找发展的突破口,未来将出现光子、量子 and 分子计算机为代表的超级计算技术革命。

4. 人工智能技术

人类的智能(artificial intelligence,AI)分别由大脑左、右两半部分管理,左脑主要负责逻辑处理,右脑是模糊处理。计算机科学中人工智能研究的是只有人的智能才具有的计算机程序或计算机系统的性能特征,研究对象是体现人的智能的程序或系统性能。它逐步发展到学习、推理、自动编程、机器人、机器视觉、专家系统以及自然语言理解等,其中机器视觉和专家系统又称为知识工程。最终,真正比较复杂的、高级的基于智能知识的系统(IKBS)应当完全是利用知识而不是狭义的实用主义的基于规则的专家系统,也就是需要自主“联想”。人们终将在分子水平和细胞水平上解析人类的高级神经活动(如感觉、认知和思维)。现阶段在图像处理领域,尤其是在较高的解释层次上正越来越多地应用 AI 技术。

1.4 图像技术的应用

除了上述三个前沿技术领域的进步外,推动图像科学发展的直接动力是各领域应用的推动。一些国家的空间规划,例如 1969 年 7 月美国 Apollo 飞船登月的电视报道需要包括图像编码、传输、校正和复原技术,在太空恶劣环境中捕获图像与传输信号也对图像

系统提出了更为苛刻的要求；超级大国对于军事优势的无休止的追求推动了在图像编码、复原、目标识别、测距和场景理解、特别是扩展不可见电磁波波谱方面应用的进展；此外，大量的从医学、核物理学、工业、农业、遥感、天文学等领域产生的图像分析的要求，特别是目前网络流媒体的发展都是图像科学发展的无穷尽的推动力。总之，略去各种复杂的应用背景，根据图像处理的最终要求和最终要求的输出形式以及所采用的主要处理技术等特点，图像技术应用领域可以粗略划分为四个方面（见图 1.2）。

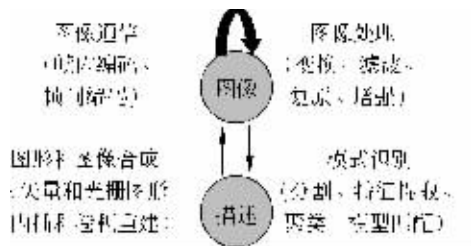


图 1.2 图像应用领域的技术特点

1. 电视与图像通信

这是最早发展起来的图像应用之一，先是在广播上（称为广播电视），以后推广到各领域（统称为工业电视或应用电视）。20 世纪 70 年代以前是模拟电视，1972 年英国独立广播公司（IBA）研制出第一台数字式电视制式转换器，标志进入数字电视的时代。随之而来的是大量的数字设备：时基校正、帧同步机、噪声抑制、数字特技等，近来的高清晰度电视，特别是由电视、通信和计算机三个领域相互融合产生了多媒体技术。在这些应用中图像处理的主要目的是以改善图像质量和压缩编码为主，例如为消除图像空间冗余度采用的帧内编码和消除时间冗余度采用的帧间编码等，整个处理过程始终保持从图像到图像的输出形式。

2. 图像处理

现在图像处理(image processing)一般是指数字图像处理或计算机图像处理。它是图像学科中最为基础的部分，图像传输、识别与理解以及图像重建的处理过程中也都离不开图像处理，在那里它被称为预处理。图像处理是以二维空间域上的线性系统理论为基础，在采用的基本技术中：变换、滤波、复原、增强等始终是保存图像特性，在不损害测量处理的意义下加以模拟和分析。目前的研究热点是采用更加高级的处理技术，例如应用二维随机过程来处理图像、应用统计判决理论降低噪声的干扰等。另外，在高速处理方面正在不断发展硬件的并行处理。一个直接应用图像处理的例子是档案工作中的模糊图像复原。

3. 图像分析

图像分析(image analysis)或称图像识别/理解(image recognition/understanding)是以模式识别(pattern recognition)为目的，其中的图像处理始终是预处理。模式识别重视的是图像的形状信息，要根据投影几何法或曲线特征分析将图像变换到抽象的描述（可以是一组数据、一串符号、一个图表等），再进一步进行分类。处理的基本过程为：分割（选

出有形状意义的部分)、提取轮廓(外边缘连接部分)、表示(求出提供几何学测量的参数,用抽象的符号加以描述)、建模(建立适合给定图像和处理问题的模型)、匹配识别(应用相关匹配法将图像区分成若干预先确定的类别)。主要应用有文字识别、医学影像诊断等。而且因为图像理解有其明确的语义学含义,常常要求关于特定问题范畴的复杂和多维的知识,所以是长期的目标。图像识别在 AI 中占有重要的地位,也称为计算机视觉(computer vision)。

4. 计算机图形与图像重建

计算机图形学(computer graphics)的主要对象是计算数据的复杂图形表示。计算机与图形的结合来源于对友好的人机交互界面的追求,20 世纪 50 年代 MIT 实现了对话式图形显示,20 世纪 60 年代中期 MIT、通用汽车公司、贝尔电话实验室和洛克希德飞机公司都进行了大规模计算机图形显示计划,到 20 世纪 70 年代结出硕果。计算机图形可分为两个基本过程:在某种坐标空间中对物体施行几何变换,从而将物体映射到屏幕上的显示坐标空间。再对被描述的物体加以光色效应或产生视觉表示(例如加纹理等)。计算机图形最早是二维,现在发展到三维。它又分成两类:以描绘产生线状图的三维透视图,被称为矢量图形(vector graphics);能产生中间色调和带颜色图像的光栅图形(raster graphics)。计算机图形学的最终目的是产生逼真的动态图像,主要应用有 CAD、交互式训练模拟器、计算机美术和动画等。

与图像分析相反,图像重建(image reconstruction)是根据一组穿过物体横截面的投影数据来重建物体的图像。影响最大的应用是在影像医学上,例如计算机断层扫描仪 CT(X 光)、ECT(γ 射线)、UCT(超声)、NMRI(核磁共振成像),另外遥感与遥测方面的 SAR(合成孔径雷达)等也是重要的应用。图像重建的主要步骤:数据采集、滤波、反投影。重建的主要算法有傅里叶空间内插法和卷积重建法(将投影数据与空间频率滤波器做卷积,然后反投影)、迭代松弛法(将图像映射为有限维向量,使用级数展开等)。

在图像科学技术的漫长的发展过程中,从光学到电子、再到光电的融合,从摄像到显像,从有线到无线传输,从器件到系统与应用,涌现出不少发明创造,极大地影响了人类社会,也获得了许多诺贝尔奖(Nobel Prize)。下面仅列举一些与图像技术发展有关的获奖情况。

1901 年第一届诺贝尔奖的物理学奖授予 Wilhelm Conrad Rontgen(德),表彰他发现 X 射线以及对 X 射线性质的研究,这些成果奠定了医学影像学的基础。

1909 年物理学奖授予 Guglielmo Marconi(意)和 Karl Ferdinand Braun(德),表彰他们对无线电通信发展的贡献,其中有 Braun 发明的 CRT 显示器件,1931 年 Zworykin 在此基础上成功地研制了电视显像管。

1921 年物理学奖授予 Albert Einstein(瑞士),表彰他对光电效应定律的贡献。在此理论的基础上,1923 年 Zworykin 发明了光电摄像管。